Technical Note: Routing Chaos versus Atlas Slot-Machine Effects in Deep Neural Networks

Motivation

In sparse or conditionally-computed neural networks, each input induces a computational route through the network.

Different routing mechanisms may exhibit fundamentally different stability properties.

Atlas-Based Routing

A globally-routed architecture determines all routing decisions at the outset:

input
→ global routing atlas
→ entire route fixed
→ computation

The route is selected once and remains fixed throughout the forward pass.

However, such systems may exhibit a **slot-machine effect**:

small input perturbation
→ different atlas code
→ different parameter subset
→ different effective mapping

If nearby inputs activate substantially different charts, generalization may degrade.

The severity of this effect depends on:

- locality preservation of the atlas;
- overlap between neighboring charts;
- smoothness of the resulting effective mappings.

Recursive Routing

Conventional deep networks determine routing recursively:

input
→ layer 1 routing
→ layer 2 routing
→ layer 3 routing
→ ...

Because later routing decisions depend on earlier activations, small perturbations may compound with depth:

small input perturbation
→ altered early routing
→ altered activations
→ additional routing changes
→ further activation changes
→ ...

This process may produce **routing chaos**.

Consequently, nearby inputs may traverse increasingly divergent computational pathways.

Competing Stability Hypothesis

Atlas-based and recursively-routed networks may therefore possess distinct failure modes:

Architecture	**Potential instability**
Global Atlas	Slot-machine chart changes
Recursive Routing	Compounding routing chaos

The key research question becomes:

> Which instability mechanism is less damaging to learning and generalization?

A globally-routed architecture may suffer occasional chart transitions but avoid compounding errors.

Conversely, recursively-routed architectures may exhibit progressively amplified routing divergence through depth.

Experimental Test

For an input x, generate

$$x' = x + \varepsilon$$

and measure route similarity throughout the network.

Metrics:

- route overlap;
- active parameter overlap;
- effective matrix similarity;
- output sensitivity;
- routing divergence as a function of depth.

Rapid decay of route similarity with depth would provide evidence for routing chaos.